Surname		Other Names	
Centre Number			Candidate Number									
Candidate Signature												

Leave blank

General Certificate of Secondary Education
June 2006

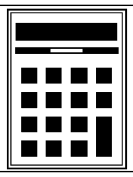
STATISTICS
Higher Tier

3311/H

H

Thursday 22 June 2006 9.00 am to 11.30 am

For this paper you must have:
- a calculator
- mathematical instruments.



Time allowed: 2 hours 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Draw diagrams in pencil.
- Fill in the boxes at the top of this page.
- Answer **all** questions in the spaces provided.
- Do all rough work in this book.

Information
- The maximum mark for this paper is 120.
- Mark allocations are shown in brackets.
- Additional answer paper and graph paper will be issued on request and must be tagged securely to this answer book.
- You are expected to use a calculator where appropriate.

Advice
- In all calculations, show clearly how you work out your answer.

For Examiner's Use	
Pages	Mark
4–5	
6–7	
8–9	
10–11	
12–13	
14–15	
16–17	
18–19	
20–21	
22–23	
24–25	
26–27	
TOTAL	
Examiner's Initials	

You may need to use the following formulae:

Mean of a frequency distribution $= \dfrac{\sum fx}{\sum f}$

Mean of a grouped frequency distribution $= \dfrac{\sum fx}{\sum f}$, where x is the mid-interval value.

Standard deviation for a set of numbers $x_1, x_2, \ldots, x_n$ having a mean value of $\bar{x}$ is given by

$$\sqrt{\frac{\sum (x - \bar{x})^2}{n}} \text{ or } \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

Standard deviation for a frequency distribution

$$\sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}} \text{ or } \sqrt{\frac{\sum fx^2}{\sum f} - \bar{x}^2}$$

The same formula applies to the standard deviation of a grouped frequency distribution where x is the mid-interval value.

Spearman's rank correlation coefficient $= 1 - \dfrac{6\sum d^2}{n(n^2 - 1)}$

Turn over for the first question

LEAVE MARGIN BLANK

Answer **all** questions in the spaces provided.

1 Nicky conducted a survey about how accurately boys measured a piece of string.
The piece of string was measured to the nearest mm.

The results of the survey are given in the stem and leaf diagram.

7	4	5	5	7	8									
8	0	0	0	0	1	2	3	4	5	Key:	8	0	means 80 mm	

Nicky then drew a second stem and leaf diagram.

7	4													
7	5	5	7	8										
8	0	0	0	0	1	2	3	4						
8	5									Key:	8	0	means 80 mm	

(a) What is the advantage of using the second stem and leaf diagram rather than the first?

.....

.....

(1 mark)

(b) Write down the shortest measurement recorded in the survey.

Answer ..... mm *(1 mark)*

(c) James measured the string as 79 mm.

Add James' measurement to the **second** stem and leaf diagram. *(1 mark)*

2 A company proposes to build a large wind turbine close to a village.
Ben designs a questionnaire to obtain opinions on the proposal from the villagers.
One of his questions is

Do you agree that the wind turbine will be a disaster for our village?

Yes, definitely ☐ No ☐

(a) Give **two** distinct criticisms of Ben's question.

Criticism 1 .....

.....

Criticism 2 .....

.....

(2 marks)

(b) Rewrite Ben's question to make it more appropriate.

.....

.....

.....

.....

(2 marks)

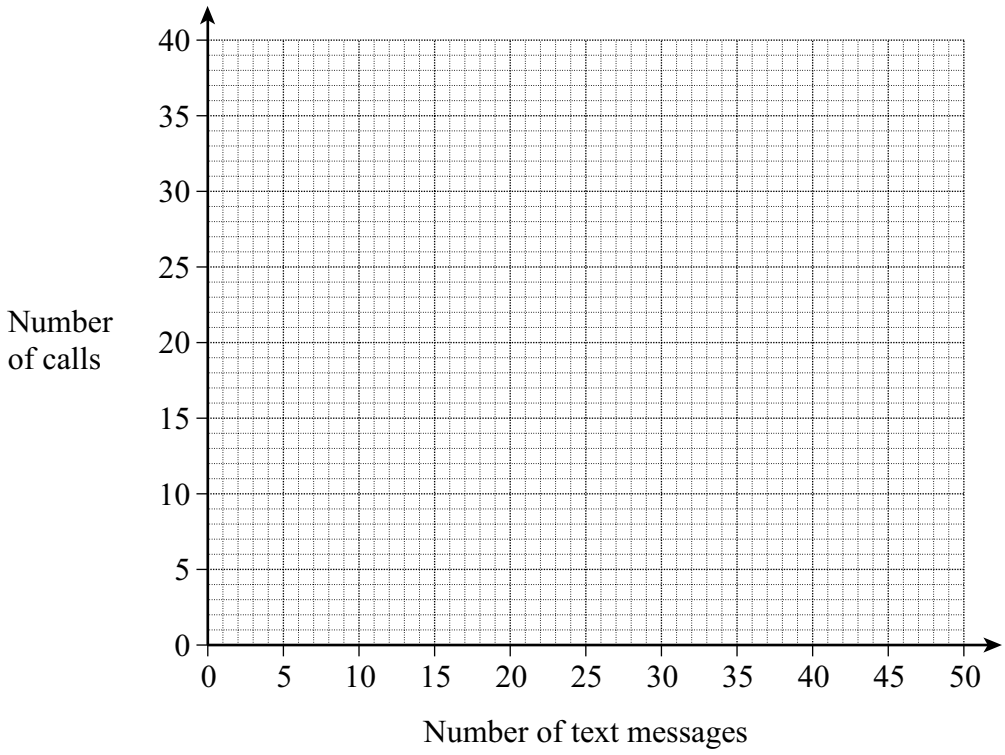
Turn over for the next question

7

3 Kelly asked 8 friends to record the number of text messages they sent and the number of calls they made on their mobile phone during one week.
The results are recorded in the table.

Number of text messages	18	22	24	25	28	32	32	35
Number of calls	36	24	26	20	13	8	10	7

(a) Use the grid to draw a scatter diagram to represent these data. *(2 marks)*



(b) Describe the correlation shown by the scatter diagram.

Answer ..... *(1 mark)*

(c) The mean number of text messages sent is 27.
What is the mean number of calls made?

.....

.....

Answer ..... *(2 marks)*

(d) Draw a line of best fit on your diagram.

(2 marks)

(e) Susan, another of Kelly's friends, sent 30 text messages during that week.

Use your line of best fit to estimate the number of calls that she made.

Answer ..... *(1 mark)*

(f) Ailsa, another of Kelly's friends, sent 45 text messages during that week.

Give **one** reason why you should not use this line of best fit to estimate the number of calls that Ailsa made.

.....

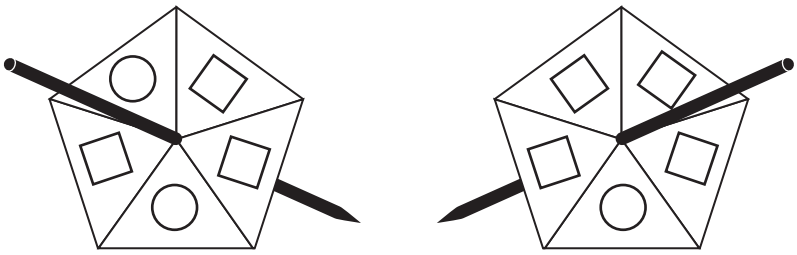
.....

.....

(1 mark)

Turn over for the next question

4 Alfie has two unbiased five-sided spinners, one red and one blue.



Red Blue

On the red spinner there are three squares and two circles.
On the blue spinner there are four squares and one circle.

Alfie spins one of the spinners. It lands on a circle.

(a) Alfie said the probability of the spinner landing on a circle was 0.2

Was the spinner red or blue? Give a reason for your answer.

Colour .....

Reason

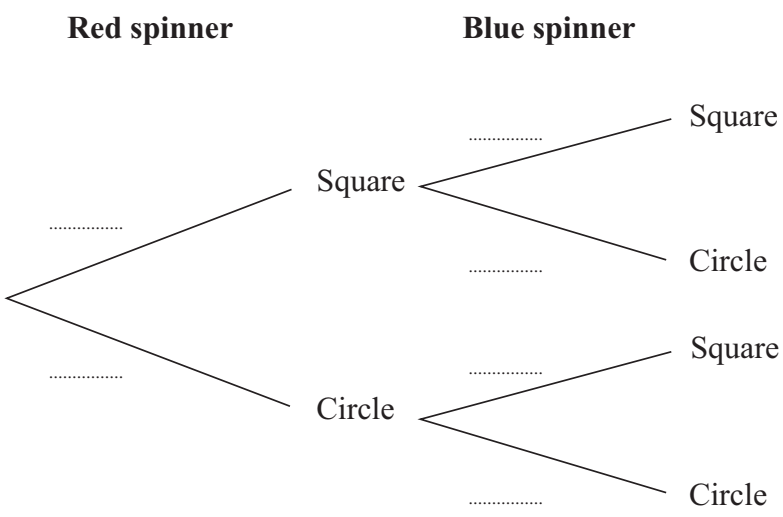
.....

.....

(1 mark)

(b) Alfie spins each spinner once.

Complete the tree diagram to show the probabilities when each spinner is spun.



(3 marks)

(c) What is the probability that both spinners land on squares?

.....

.....

Answer ..... *(2 marks)*

Turn over for the next question

6

5 The table shows the percentage by age for each ethnic group of the UK population 2001 – 2002.

Ethnic Group \ Age	Under 16	16 – 34	35 – 64	65 and over
White	19	25	40	16
Mixed	55	27	16	2
Indian	22	34	38	6
Pakistani	35	36	25	4
Bangladeshi	38	38	21	3
Other Asian	22	36	38	4
Black Caribbean	24	25	42	9
Black African	33	35	30	2
Other Black	35	34	26	5
Chinese	20	40	35	5
Other	20	37	39	4

Source: *Adapted from Office for National Statistics, Summer 2003*

(a) Which ethnic group had the largest percentage of its population under 16 years of age?

Answer ..... *(1 mark)*

(b) What was the difference between the percentages of Chinese ethnic group and Black African ethnic group aged 35 – 64 years?

.....

Answer ..... *(2 marks)*

(c) Give one similarity and one difference between the age profiles of the White ethnic group and the Indian ethnic group.

Similarity .....

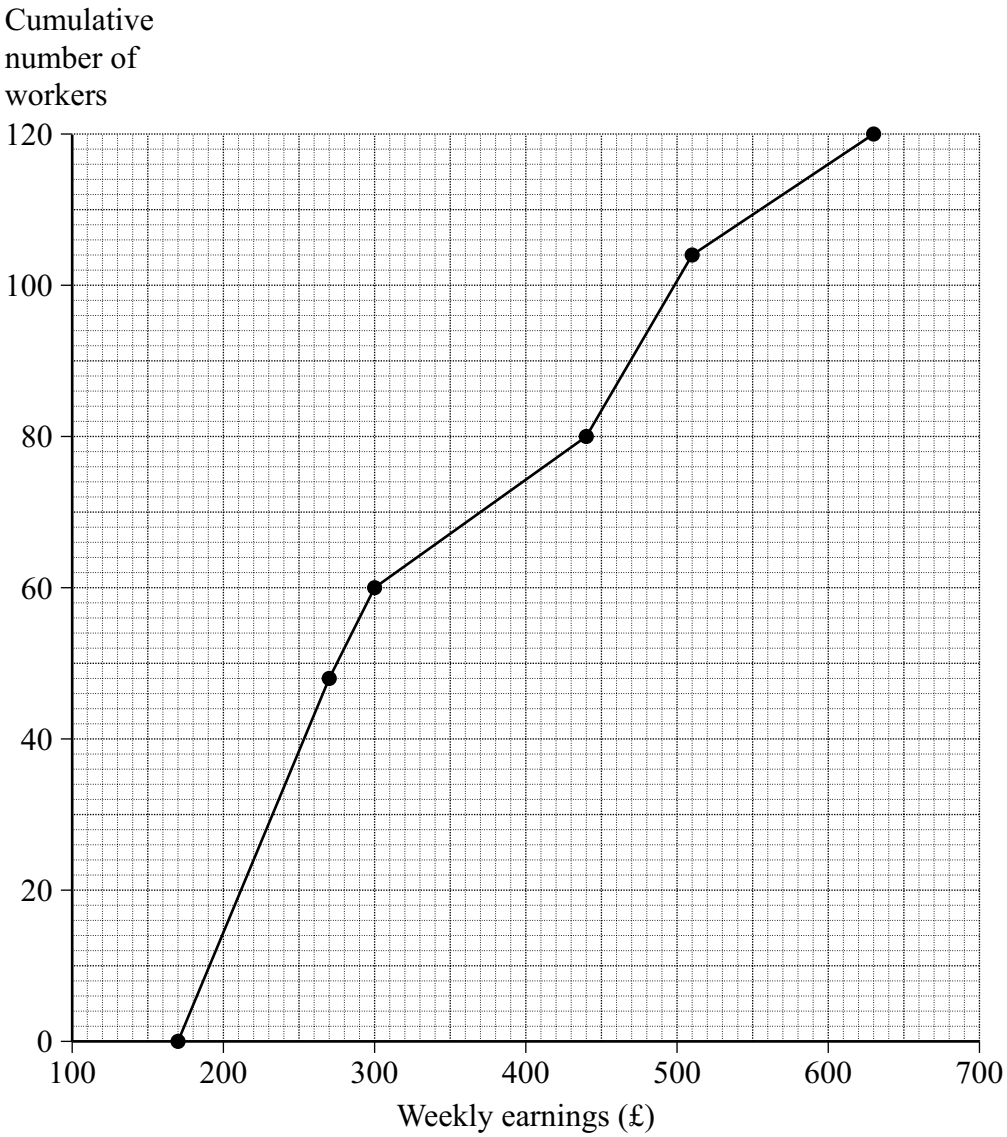
.....

Difference .....

.....

(2 marks)

6 The cumulative frequency polygon shows the distribution of weekly earnings of a sample of 120 male manual workers in the ceramics industry.



(a) Use the graph to estimate

(i) the median

Answer £ ..... *(1 mark)*

(ii) the interquartile range

.....

.....

Answer £ ..... *(2 marks)*

Question 6 continues on the next page

8

(iii) the percentage of workers earning under £320 per week

.....

.....

.....

Answer ..... % *(3 marks)*

(iv) the 9th decile.

.....

.....

.....

Answer £ ..... *(2 marks)*

(b) The following information was found from a sample of 120 female manual workers in the ceramics industry.

The median of the weekly earnings was £230.

25% of the sample had weekly earnings more than £280.

The interquartile range was £100.

No one earned less than £120 per week or more than £420 per week.

Six workers earned more than £390 per week.

(i) Use your answers to part (a) and the information on female earnings to make **two** statements that support the following hypothesis:

'Female workers in the ceramics industry have lower **and** less variable weekly earnings than male workers in the ceramics industry.'

Statement 1 .....

.....

Statement 2 .....

.....

(2 marks)

(ii) Describe another source of data that could be used to explore this hypothesis.

.....

.....

(1 mark)

(c) One method of calculating a measure of skewness is:

$$\frac{\textbf{Upper Quartile + Lower Quartile – 2 (median)}}{\textbf{Upper Quartile – Lower Quartile}}$$

(i) Use this method to calculate separate measures of skewness for the weekly earnings of male **and** female workers.

.....

.....

.....

.....

.....

.....

.....

Answer Male workers .....

Female workers ..... *(3 marks)*

(ii) Use your answers to part (c)(i) to describe the shape of each distribution.

Answer Male workers .....

Female workers ..... *(2 marks)*

Turn over for the next question

7 A local newspaper investigates unemployment in the town of Stokeham.

The table gives unemployment data for Stokeham.

Age Group	Population in thousands	Number Unemployed	Standard Population
16–24	16	1020	25%
25–44	28	1540	30%
45–54	32	1206	30%
55–64	14	680	15%

(a) Calculate the crude unemployment rate for Stokeham.

.....

.....

.....

.....

.....

.....

Answer ..... *(3 marks)*

(b) Find the standardised unemployment rate for Stokeham.

.....

.....

.....

.....

.....

Answer ..... *(4 marks)*

(c) What is the advantage in using the standardised rate rather than the crude unemployment rate in this case?

.....

.....

(1 mark)

(d) The nearby town of Bloomfield has a standardised unemployment rate of 71.4 per thousand.

Michelle is moving into one of these towns.
In which town would she stand a better chance of finding employment?

Give a reason for your answer.

Town .....

Reason .....

.....

.....

(1 mark)

Turn over for the next question

9

8 Records for a local library show for each book whether it is in the fiction, non fiction or classics category and whether it is a hard back or soft back version.

When the library closed on Wednesday last week 2700 books were out on loan.

Of the books on loan 72% were in the fiction category.

Of the 620 hard back books on loan 55% were in the non fiction category and 25% in the classics category.

In total 176 classics books were on loan.

(a) Complete the table, entering the number of books on loan in each case.

Version / **Category**	Hard back	Soft back	Totals
Fiction			
Non Fiction			
Classics			176
Totals	620		2700

(4 marks)

(b) A library record for a book on loan is chosen at random.

Use the table to calculate the probability that the book is

(i) non fiction and a soft back version

.....

Answer ..... *(1 mark)*

(ii) non fiction or a hard back version

.....

Answer ..... *(2 marks)*

(iii) fiction, given that it is a soft back version.

.....

Answer ..... *(2 marks)*

(c) How many of the first 200 books taken out on loan on the following day would you expect to be hard back classics?

.....

.....

.....

Answer ..... *(2 marks)*

Turn over for the next question

9 Alan works for the home delivery service of a large supermarket chain.

He has been asked by his manager to find out how long it takes to select and pack a sample of customer orders.

The table shows the results based on a sample of 75 orders.

Table 1

Time, t (minutes)	Number of orders
$0 \leqslant t < 3$	1
$3 \leqslant t < 6$	5
$6 \leqslant t < 9$	7
$9 \leqslant t < 12$	24
$12 \leqslant t < 15$	24
$15 \leqslant t < 18$	8
$18 \leqslant t < 21$	6

To simplify the data Alan regrouped the 75 values as follows

Table 2

Time, t (minutes)	Number of orders
$0 \leqslant t < 6$	6
$6 \leqslant t < 10$	8
$10 \leqslant t < 11$	22
$11 \leqslant t < 14$	18
$14 \leqslant t < 21$	21

(a) Give two reasons why the first table is a better form of grouping than the second.

Reason 1 .....

.....

Reason 2 .....

.....

(2 marks)

(b) State the modal class(es) for the distributions in Tables 1 and 2.

Answer Table 1 .....

Table 2 ..... *(2 marks)*

(c) Use the data in **Table 2** to draw a histogram on the grid below.

.....

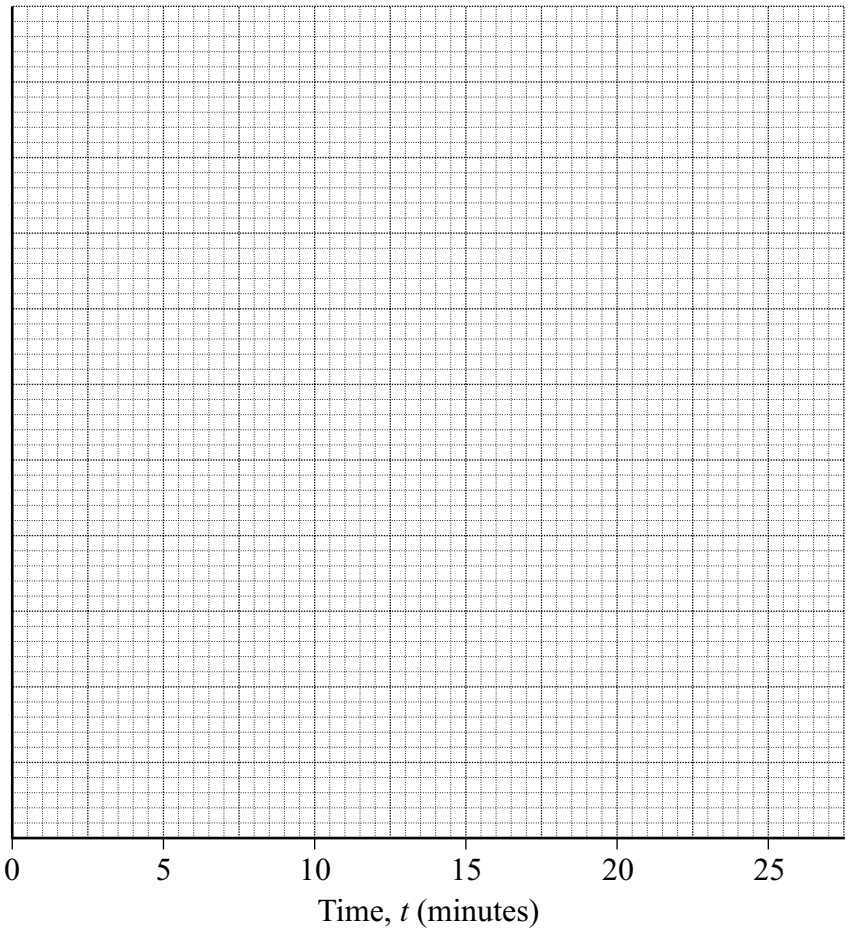
.....

.....

.....

.....

(5 marks)



(d) Records show that 10% of all orders packed by supermarket staff have missing items.

Alan selects 5 packed orders at random.
He checks each order for missing items.
Each order packed is independent of all previous orders.

Calculate the probability that none of these orders have missing items.

.....

.....

.....

.....

Answer ..... *(3 marks)*

12

10 The diagram shows a factory layout divided into four different work areas.

The numbers of male and female staff in each area are also given.

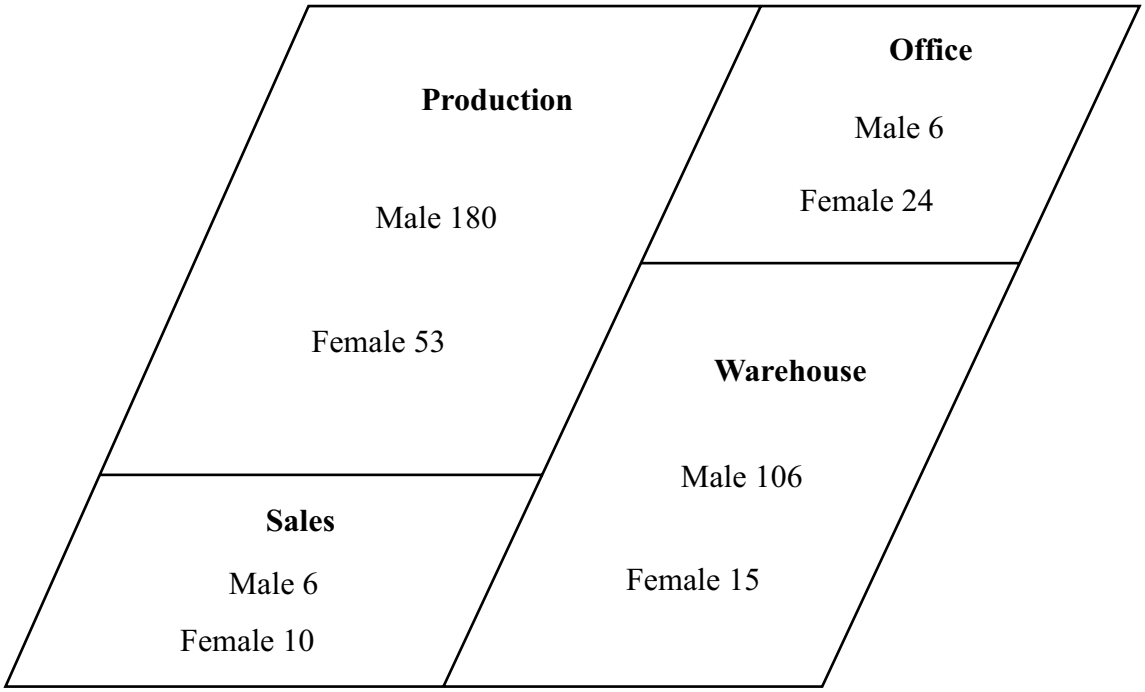


Diagram not to scale

The factory owner wishes to undertake a survey to find the reaction of the staff to the introduction of a new bonus payment scheme.

He decides to take a systematic sample of 20 male production staff.

(a) Explain how this sample could be selected.

.....

.....

.....

.....

(3 marks)

(b) Give **two** reasons why this sample would be unrepresentative of the whole staff.

Reason 1 .....

.....

Reason 2 .....

.....

(2 marks)

(c) As an alternative the owner is advised to take a sample, stratified by sex and work area, of 50 of the 400 staff.

(i) Calculate the number of sales staff to be included in the sample.

.....

.....

Answer ..... *(2 marks)*

(ii) Calculate the number of female office staff to be included in the sample.

.....

.....

Answer ..... *(2 marks)*

(d) Part of the survey will involve interviewing staff to find their views on plans to change the number of hours worked each week in the factory.

One of the questions to be asked will be

> *What is your opinion of the proposal to increase the number of hours worked each week to 39?*

Describe two types of scale that could be used to measure the opinions in this case.

Type 1 .....

.....

.....

.....

Type 2 .....

.....

.....

.....

(4 marks)

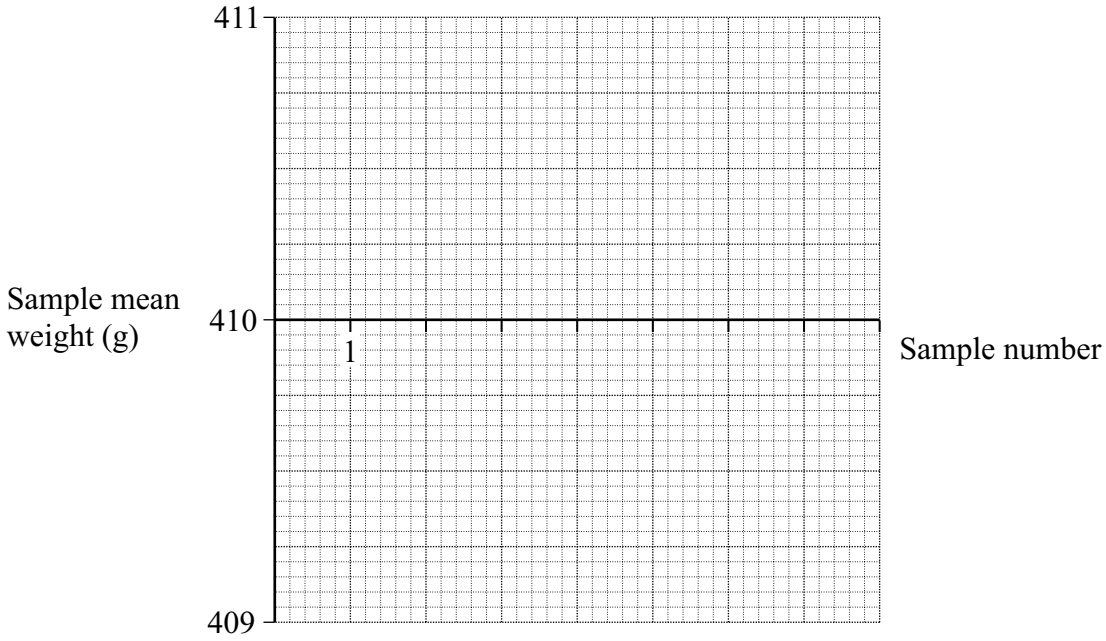
Question 10 continues on the next page

(e) The factory uses one machine to produce tins of food.
The Quality Control staff check the process by taking samples of 5 tins every half hour and recording the mean weight of each sample.

The target value for the mean weight is 410 g
The table shows the results of the first 8 samples taken from the machine.

Sample number	Sample mean weight (g)
1	410.30
2	409.65
3	410.25
4	410.30
5	410.50
6	410.95
7	410.30
8	409.70

(i) On the grid below draw a chart to show this data.



(2 marks)

(ii) At which point was the process stopped and the machine adjusted? Justify your choice.

.....

.....

.....

(2 marks)

(f) The machine has an automatic counter that records the total number of tins produced every hour.

(i) What form of data collection procedure is used in this case?

Answer ..... *(1 mark)*

(ii) Give one advantage in using this method.

Advantage .....

.....

(1 mark)

Turn over for the next question

6

11 Joan is a Road Safety Officer for a City Council.

Part of her work involves recording the number of vehicles exceeding the speed limit as they pass local schools.

The following table gives the data recorded over a 120 day period e.g. on 42 days two vehicles per day exceeded the speed limit.

Number of vehicles per day exceeding speed limit	2	5	6	10	14	15	20
Number of days	42	28	18	14	10	5	3

(a) Calculate the mean and standard deviation of the number of vehicles per day exceeding the speed limit.

Give your answers to two decimal places.

.....

.....

.....

.....

.....

Answer mean .....

standard deviation ..... *(5 marks)*

(b) Due to a fault on the recording equipment, Joan's records did not show a further two vehicles **each** day which were exceeding the speed limit.

What effect will this error have on the values for

(i) the mean

Answer .....

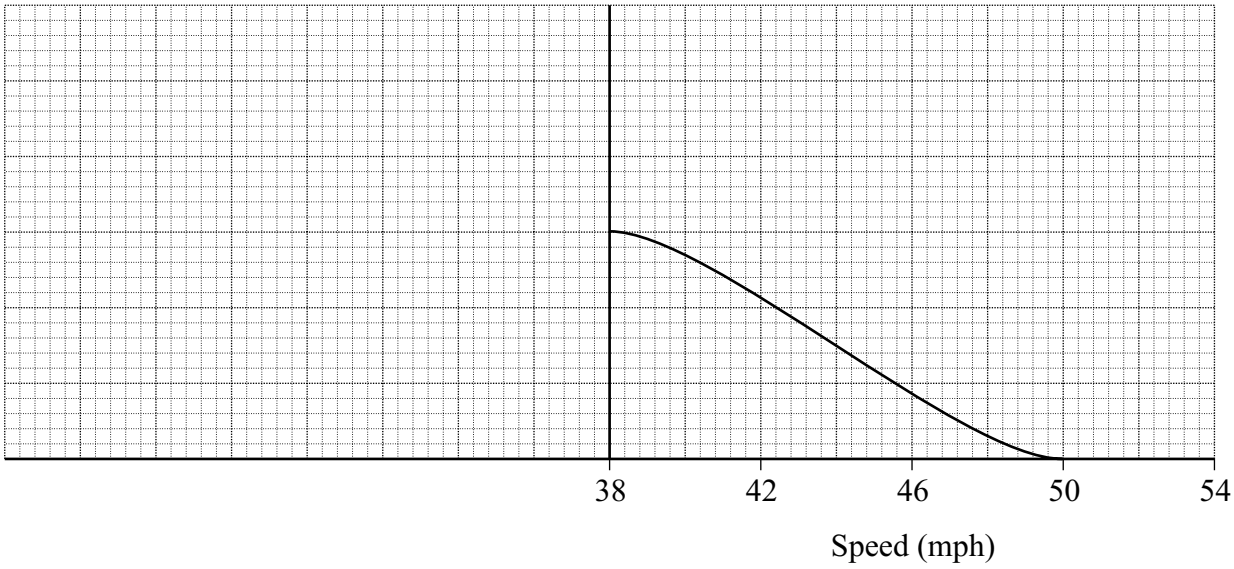
(ii) the standard deviation?

Answer ..... *(2 marks)*

(c) During the same period of time two other Road Safety Officers recorded equal numbers of actual traffic speeds at different locations in the city.

The speeds recorded by the first Road Safety Officer are normally distributed.

(i) On the grid below complete the diagram for this distribution.



(2 marks)

(ii) The speeds recorded by the second Road Safety Officer were also normally distributed.
They had a mean of 44 mph and standard deviation of 2 mph

On the same grid draw a diagram to represent this distribution.

(3 marks)

(iii) For the first Road Safety Officer, what proportion of his records will show speeds of 38 mph or less?

Answer ..... *(1 mark)*

(iv) For the second Road Safety Officer, what proportion of his records will show speeds above 52 mph?

Answer ..... *(1 mark)*

12 An expert from the local antiques club agreed to challenge a number of contestants to correctly rank eight items of Victorian furniture according to their value.

John agreed to take part and his **rankings** along with those of the expert were as follows

Exhibit	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Expert	1	3	6	7	8	2	5	4
John	8	5	2.5	4	1	6	2.5	7

(a) Calculate the value of Spearman's rank correlation coefficient for the two sets of data.

.....

.....

.....

.....

.....

Answer ..... *(4 marks)*

(b) Interpret, in context, your answer to part (a).

.....

.....

(1 mark)

(c) A further eight contestants entered the competition.
The values of the correlation coefficients were

0.35 −0.43 0.71 0.05 −0.36 −0.02 0.92 −0.81

(i) Which two of these values show that there is almost no correlation between the individual rankings of that contestant and those of the expert?

Answer ..... *(2 marks)*

(ii) Which of these values shows the strongest correlation between the expert and contestant?

Answer ..... *(1 mark)*

(d) John was asked to draw two comparative pie charts to show the value of UK exports of Victorian furniture to Europe in 1994 and 2004.

A pie chart of radius 3 cm was drawn to show the 1994 export figure of £107 000. The export figure for 2004 was £196 400.

Calculate the radius of the pie chart used for the 2004 total.

Give your answer to two decimal places.

.....

.....

.....

.....

Answer ..... cm *(4 marks)*

END OF QUESTIONS

12

There are no questions printed on this page